

Vorwort

Das vorliegende Buch basiert auf dem 1996 erschienenen Werk *Kognitive Modellierung: Eine Einführung in logische und algorithmische Grundlagen* von Ute Schmid und Martin Kindsmüller. Ziel des Buches war es, Studierenden und Forschenden aus dem Bereich Kognitionspsychologie eine Einführung in zentrale Methoden der Künstlichen Intelligenz zu geben. Im Gegensatz zu Standardlehrbüchern zur Künstlichen Intelligenz werden die Themen so eingeführt, dass keine Vorkenntnisse im Bereich Informatik notwendig sind. Alle Themen werden zudem mit expliziten Bezügen zur kognitionswissenschaftlichen Forschung motiviert, und es wird besonderes Augenmerk auf viele klassische Grundlagenartikel gelegt, ohne die die aktuelle Entwicklung kaum verständlich wäre. Das Buch erschien im letzten der sogenannten KI-Winter – also in einer Zeit, in der kaum Interesse an Künstlicher Intelligenz (KI) außerhalb der engeren Forschungsgemeinschaft bestand. Dies hat sich mit dem Aufkommen neuer Ansätze, insbesondere *Deep Learning* und *Generative KI*, schlagartig geändert. Das Thema KI erfährt eine nie dagewesene Aufmerksamkeit und wird in nahezu allen wissenschaftlichen Disziplinen diskutiert, so auch in der Psychologie. Zunehmend werden Methoden der KI in vielen Bereichen angewendet – in der Psychologie als Ergänzung zu statistischen Methoden, zur Modellierung kognitiver Prozesse oder als Untersuchungsgegenstand.

Entsprechend war es an der Zeit, das oben genannte Buch in gründlich überarbeiteter und deutlich aktualisierter Form neu aufzulegen. Anders als 1996 ist die Bezeichnung „Künstliche Intelligenz“ nun wieder salonfähig und muss nicht hinter der Bezeichnung „Kognitive Modellierung“ versteckt werden. Die Initiative für das vorliegende Buch *Kognitive Künstliche Intelligenz* kam von Marco Ragni, der die Überarbeitung des ursprünglichen Werkes, bei dem Ute Schmid Erstautorin war, maßgeblich umgesetzt hat. Beide Autoren hoffen, dass dieses Buch Studierenden, Forschenden sowie Praktikerinnen und Praktikern einen hilfreichen Einstieg in Themen und Methoden der Künstlichen Intelligenz bietet.

Wir bedanken uns bei Christopher von Bülow für die kritische Durchsicht und finale Editierung des Manuskripts. Dominik Bär danken wir herzlich für die vielfältige Unterstützung, einschließlich der Bildbearbeitung. Ebenso sind wir Sara Todorovikj und Milena Stella Jans für ihre Unterstützung bei der Bildbearbeitung verbunden. Unser Dank geht besonders auch an unsere Lektorin beim Springer-Verlag, Marion Krämer, für ihre Geduld und Unterstützung.

Marco Ragni

Ute Schmid

Juli 2025

Zusätzliche Online-Materialien inkl. Dozentenfoliensatz zu diesem Buch finden Sie auf: ► <https://www.lehrbuch-psychologie.springernature.com>



Einführung

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer-Verlag GmbH, DE, ein Teil von Springer Nature 2025

M. Ragni, U. Schmid, *Kognitive Künstliche Intelligenz*, https://doi.org/10.1007/978-3-662-69498-5_1

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer-Verlag GmbH, DE, ein Teil von Springer Nature 2025
Aus: Ragni, M. & Schmid, U. (2025). *Kognitive Künstliche Intelligenz*. Springer.

Kognitive Künstliche Intelligenz (KKI) ist eine junge interdisziplinäre Wissenschaft, die sich mit natürlichen wie künstlichen, konkreten wie abstrakten kognitiven Prozessen auseinandersetzt und prädiktive¹ sowie erklärbare algorithmische Modelle² für diese entwickelt.

Sie stellt die formale Nachbardisziplin zur Kognitionswissenschaft dar, die ihrerseits von Kognitionspsychologie, Neurowissenschaft, Linguistik und analytischer Philosophie inspiriert ist. Kognition beim Menschen bezeichnet dabei alle mentalen, d. h. geistigen, Prozesse, die für die Verarbeitung von Information relevant sind. Dies umfasst deren Erwerb (durch Wahrnehmungsprozesse und Aufmerksamkeit), ihre spezifische mentale Repräsentation (z. B. im menschlichen Arbeitsgedächtnis), das Schlussfolgern (also die Verarbeitung der repräsentierten Information), das Anpassen vorhandenen Wissens (durch Lernen), die Verknüpfung von Informationen (Analogie), das Gewinnen von Einsichten bei neuen Aufgabenstellungen (Problemlösen) und die Kommunikation mit der Umwelt (z. B. Sprache; Kluwe 2000). Die Kognitionswissenschaft beschreibt menschliche kognitive Prozesse durch Modelle über menschliche Informationsverarbeitung (vgl. schon Strube 1993). Somit bildet die Kognitionswissenschaft einen wichtigen Baustein der KKI. Zugleich geht die KKI über die Kognitionswissenschaft hinaus, indem sie die kognitiven Konzepte nicht nur als Grundlage von KI betrachtet, sondern beobachtete kognitive Prozesse formalisiert und systematisch weiterentwickelt. Wesentliche Forschungsthemen beider Nachbardisziplinen sind alltägliche Ausprägungen von Intelligenz wie Wahrnehmung, Denken, Problemlösen, Sprache, Wissensorganisation und Lernen.

Kognitive Künstliche Intelligenz setzt eine interdisziplinäre Zusammenarbeit voraus, bei der gleichzeitig die einzelnen Disziplinen ihre spezifischen Schwerpunkte und Forschungsmethoden beibehalten. Die Verzahnung der von den verschiedenen Disziplinen verwendeten Methoden bringt ein mächtigeres Instrumentarium zur Erforschung kognitiver Prozesse hervor, als es jede Einzeldisziplin liefern könnte. Auch die Theoriebildung in den genannten Disziplinen ist häufig kognitiv orientiert. So sind viele kognitive Theorien ursprünglich von klassischen Ansätzen der KI beeinflusst, und die Ergebnisse kognitionspsychologischer Arbeiten beeinflussten und beeinflussen die Entwicklung der Künstliche-Intelligenz-Forschung.

Die interdisziplinäre Verankerung der KKI in Forschungsarbeiten macht sich zwangsläufig in der Lehre an der Universität bemerkbar. Beispielsweise werden bereits in einführenden Lehrbüchern zur kognitiven Psychologie (z. B. Anderson 2007b) Ansätze vorgestellt, die ihre Wurzeln in der KI haben oder von Arbeiten aus der KI beeinflusst sind. Insbesondere gilt dies für die Themen Wissensrepräsentation und Problemlösen. Der informatische Hintergrund dieser Ansätze wird zwar häufig erwähnt, die formalen Methoden, auf denen sie basieren, werden jedoch meist nicht eingeführt.

Ziel dieses Buches ist es, Prinzipien, Intuitionen, Ansätze und formale Methoden der KI zu vermitteln, die sich zur Modellierung kognitiver Prozesse als erfolgreich herausgestellt haben. Dabei haben wir uns bemüht, den Stoff so aufzubereiten, dass er auch von Lesern ohne Vorkenntnisse in Mathematik und Informatik nachvollzogen werden kann.

1 Prädiktive Modelle leiten aus vergangenen Beobachtungen Muster ab, um zukünftige Entwicklungen oder Verhalten vorherzusagen.

2 Erklärbare algorithmische Modelle (auch manchmal erklärbare KI) zielen darauf ab, die Funktionsweise komplexer KI-Modelle – insbesondere solcher, die nicht von Natur aus leicht verständlich sind (wie tiefes Lernen) – verständlich zu machen. Dadurch kann man als Mensch besser nachvollziehen, warum ein Modell bestimmte Entscheidungen trifft.

Es werden zunächst die wissenschaftstheoretischen und philosophischen Grundlagen der KKI dargestellt (► Kap. 2).

Danach ist das Buch in drei Teile gegliedert: Im ersten Teil des Buches werden formale Grundlagen für psychologische und kognitive Modellvorstellungen zur Repräsentation von deklarativem Wissen vermittelt. Hier beziehen wir uns vor allem auf den Ansatz der semantischen Netze, gehen aber auch kurz auf den schematheoretischen Ansatz ein. Beide Ansätze haben ihr Fundament in der formalen Logik. Als Grundlage für das Verständnis logischer Formalisierung führen wir zunächst wesentliche Konzepte der Mengenlehre ein (► Kap. 3). Wir erklären, wie hierarchische semantische Netze durch mengentheoretische Konzepte beschrieben werden können. In ► Kap. 4 und 5 stellen wir die Syntax der Aussagen- und der Prädikatenlogik vor und zeigen die Anwendbarkeit logischer Schlussregeln. Wir veranschaulichen dann, wie durch Überführung natürlichsprachiger Sätze in die Syntax der Aussagenlogik semantische Netze konstruiert werden können. Die wichtigste Methode zum Schlussfolgern aus gegebenem Wissen ist der Theorembe-
weis. Diesen Ansatz stellen wir in ► Kap. 6 dar. Die Prädikatenlogik ist die wesentliche Grundlage für die Programmiersprache Prolog (► Kap. 7). Wir führen Prolog beispielhaft anhand der Implementation eines semantischen Netzes ein.

Der zweite Teil des Buches dient der Vermittlung der Grundlagen für die kognitive Modellierung von Problemlösefertigkeiten. Hier stellen wir zunächst die Grundbegriffe der theoretischen Informatik dar; insbesondere führen wir Algorithmen und formale Sprachen ein (► Kap. 8). Viele Ansätze zur Modellierung menschlicher Denk- und Problemlöseprozesse basieren auf diesen Grundlagen. Im nächsten Kapitel (► Kap. 9) stellen wir dar, wie Probleme repräsentiert werden können, und zeigen, dass Problemlösen als Suche in einem „Problemraum“ beschrieben werden kann. Wir stellen zwei grundlegende Suchverfahren vor (► Kap. 10), die dann in ► Kap. 11 zu heuristischen Suchverfahren erweitert werden. Heuristische Suchverfahren sind ein zentraler Bestandteil von Produktionssystemen, die wir in ► Kap. 12 einführen. Zum Abschluss des zweiten Teils werden wir uns mit dem Lernen von Regeln (► Kap. 13) beschäftigen sowie eine kurze Einführung in den Bereich der neuronalen Netze geben (► Kap. 14).

Ziel des dritten Teils ist es, die Anwendung der vermittelten Grundlagen in verschiedenen Forschungsbereichen aufzuzeigen. Exemplarisch stellen wir die Bereiche Lernen und Expertise (► Kap. 15 und 16), Sprachverarbeitung (► Kap. 17 und 18) sowie Textverstehen (► Kap. 19 und 20) dar. Das Buch schließt mit einem kurzen Ausblick und einer Einschätzung des Potentials und der Grenzen der dargestellten Ansätze zur Modellierung kognitiver Prozesse (► Kap. 21) und greift damit die in ► Kap. 2 dargestellten Themen wieder auf.

Wir legen vor allem Wert auf eine umfassende und verständliche Einführung der formalen Grundlagen. Aus diesem Grund stellen wir nicht nur die aktuellen Ansätze vor, sondern beziehen uns wann immer möglich auf die „Klassiker“, welche oftmals viele Entwicklungen vorweggenommen haben. Am Ende jedes Kapitels geben wir Hinweise auf weiterführende Arbeiten. Deutsche Übersetzungen von Fachtermini verwenden wir immer, wenn die entsprechenden Begriffe in der deutschsprachigen Literatur gebräuchlich sind. Wo dies nicht der Fall ist, geben wir den englischen Begriffen den Vorzug, auch wenn dadurch unschöne Kombinationen aus englischen und deutschen Worten entstehen. Wir fokussieren auf die Darstellung symbolischer Ansätze, gehen aber auch auf die Prinzipien der Modellierung mit neuronalen Netzen ein.

Wir gehen in mehreren Kapiteln auf die Umsetzung der dargestellten Formalismen in Computerprogrammen ein. Da es uns nicht als sinnvoll erscheint, die Einführung formaler

Konzepte durch die Erläuterung technischer Begriffe zu unterbrechen, führen wir die Definitionen informatischer Konzepte in einem Glossar auf.

Unser Anliegen ist, dass nach Lektüre dieses Buches Spezialliteratur zur kognitiven Modellierung sowie Lehrbücher zur Künstlichen Intelligenz ohne große Verständnisprobleme gelesen werden können. Das Buch ist so aufgebaut, dass die in Teil I und II dargestellten Inhalte in einer einsemestrigen Lehrveranstaltung vermittelt werden können. Die in Teil III dargestellten Themen sollen Anregungen für die Auseinandersetzung mit speziellen Forschungsgebieten der Kognitionswissenschaft geben. Hier empfiehlt es sich unserer Meinung nach, spezielle Veranstaltungen zu konzipieren, die über die von uns dargestellten Inhalte hinausgehen.



Was ist Kognitive Künstliche Intelligenz?

Inhaltsverzeichnis

- 2.1 Der Informationsverarbeitungsansatz – 8
- 2.2 Philosophische Grundlagen der Kognitionsforschung – 10
- 2.3 Die Methode der kognitiven Modellierung – 13
- 2.4 Zur Vertiefung – 15

„Kognitive Künstliche Intelligenz“ ist ein Sammelbegriff, unter dem die Forschung zu kognitiven Strukturen, Prozessen und wissensbasierten technischen Systemen zusammengefasst wird. Klassische Forschungsthemen sind Wahrnehmung, Denken, Problemlösen, Sprache, Wissensorganisation und Lernen. Disziplinen, in denen Fragestellungen aus diesen Bereichen bearbeitet werden, sind insbesondere Künstliche Intelligenz, kognitive Psychologie, Neurowissenschaften, Linguistik und Philosophie.

Zentrales Forschungsanliegen der Künstlichen Intelligenz ist der Entwurf von Algorithmen, die komplexe Probleme aus den oben genannten Bereichen bewältigen können. In der kognitiven Psychologie werden Modelle menschlicher Kognition aufgestellt und empirisch überprüft. Die Neurowissenschaften beschäftigen sich mit Aufbau und Funktionsweise von Nervensystemen, insbesondere des Gehirns, und erforschen diese empirisch. Die Linguistik wiederum definiert normative und deskriptive Modelle über die Regularitäten sprachlicher Strukturen. Und in der Philosophie werden Grundlagen und Grenzen des Denkens und Verhaltens analysiert.

Kognitive Strukturen und Prozesse werden in jeder Disziplin mit deren eigenen Zielsetzungen und Methoden erforscht. Zum Teil haben sich bereits so starke interdisziplinäre Anknüpfungspunkte zwischen je zwei Disziplinen ausgebildet, dass sich Forschungsgebiete wie Neuropsychologie, Neuroinformatik, Computerlinguistik und Psycholinguistik herausgebildet haben.

In diesem Buch wird insbesondere die Interaktion von kognitiver Psychologie und Künstlicher Intelligenz behandelt. Der Schnittbereich dieser beiden Disziplinen ist die Modellierung kognitiver Prozesse. Sowohl kognitive Psychologie als auch Künstliche Intelligenz sind von ihrem Ursprung her stark interdisziplinär orientiert. Die Wende der Psychologie in den 50er-Jahren des vergangenen Jahrhunderts vom Behaviorismus hin zum Paradigma der Kognitionsforschung wurde unter anderem durch das Aufkommen der Kommunikations- und Informationstheorie (Shannon 1948) und der Kybernetik (Wiener et al. 2019) sowie von den Arbeiten des Linguisten Chomsky eingeleitet. Kommunikationstheorie, Informationstheorie und Kybernetik bildeten gleichzeitig wesentliche Grundlagen für die in den 60er-Jahren des 20. Jahrhunderts entstehende Informatik.

Als sich entwickelnde Teildisziplin der Informatik entstand die Künstliche Intelligenz. So wurde der Begriff „artificial intelligence“ (deutsch: Künstliche Intelligenz, kurz KI) erstmalig 1956 auf einer Computer-Konferenz¹ verwendet. Zu den „Gründervätern“ der KI gehörten neben Alan Turing, John McCarthy und Marvin Minsky auch Forscher wie Allen Newell und Herbert A. Simon, die die kognitive Psychologie entscheidend mitgeprägt haben.

Das gemeinsame Interesse der KI und der kognitiven Psychologie an der Modellierung menschlicher Kognition führte jedoch nicht zum Aufgehen beider Disziplinen in der Kognitionsforschung (Schmalhofer und Wetter 1988). Die kognitive Psychologie hat primär das Ziel, menschliche Informationsverarbeitungsprozesse zu erforschen. Eine wesentliche Methode, sowohl zur Bildung als auch zur Prüfung von Hypothesen, ist dabei die empirische Untersuchung der kognitiven Leistungen des Menschen. Die Künstliche Intelligenz konzentriert sich darauf, effiziente Systeme zu entwickeln, die komplexe Aufgaben bewältigen können, die ursprünglich nur durch menschliche Intelligenz lösbar waren. Ein Beispiel dafür sind Expertensysteme, die bereits in den 1970er-Jahren entwickelt wurden, um menschliche Experten durch die Abbildung ihres Wissens und

1 „The Dartmouth Summer Research Project on Artificial Intelligence“, organisiert von McCarthy. Unter den Vortragenden waren neben McCarthy unter anderem Minsky und Newell.

ihrer Schlussfolgerungsmechanismen zu simulieren. Geht man dabei von der Annahme aus, dass menschliche Informationsverarbeitung effizient ist, so ist es für die Künstliche Intelligenz von Interesse, beim Entwurf ihrer Systeme Erkenntnisse über menschliche Informationsverarbeitungsprozesse zu berücksichtigen.

Andererseits unterliegt die psychologische Theoriebildung – wie alle wissenschaftlichen Theorien – den Kriterien der Konsistenz und Prüfbarkeit (Schneewind 1977). Konsistente Theoriebildung ist jedoch nur durch eine hinreichend präzise sprachliche Darstellung der Konzepte möglich. Natürlichsprachige Darstellungen von Theorien sind oftmals unpräzise (Westmeyer 1977): Einerseits werden in verschiedenen Theorien über denselben Gegenstand unterschiedliche Begriffe verwendet, andererseits können in verschiedenen Theorien unterschiedliche Ideen mit denselben Begriffen ausgedrückt werden. Dies macht die Theorien schwer miteinander vergleichbar. Oft ist es sogar schwierig, für eine einzige Theorie zu prüfen, ob die ihr zugrunde liegenden Annahmen widerspruchsfrei sind und ob alle Annahmen, die notwendig sind, um den behandelten Gegenstand zu beschreiben, auch tatsächlich explizit formuliert wurden. So lässt sich menschliches Vergessen in der sogenannten artikulatorischen Schleife (also dem Gedächtnissystem, welches uns hilft, verbale Information wie Telefonnummern uns zu merken) auf 144 verschiedene Arten implementieren (Lewandowsky und Farrell 2011).

Mit einer formalen Darstellung von Theorien können diese Probleme ausgeräumt werden (Konerding 1992). Unter Formalisierung versteht man, dass Theorien in einer formalen Sprache mit streng festgelegter Syntax und Semantik dargestellt werden. Alle Aussagen einer Theorie könnten zum Beispiel in der Syntax der Prädikatenlogik (► Kap. 5) formuliert werden. Zu jedem eingeführten Symbol in dieser Sprache muss zudem genau festgelegt werden, was es bezeichnen soll. Gleiche Symbole bezeichnen dann immer exakt gleiche Konzepte.

In den 70er-Jahren des 20. Jahrhunderts wurde im Rahmen der mathematischen Psychologie versucht, psychologische Theorien zu formalisieren (Krantz et al. 1974). Mathematische Modelle sind jedoch in den meisten Fällen statisch. Sie sind sehr gut dazu geeignet, Beziehungen zwischen vorliegenden Stimuli und resultierendem Verhalten formal abzubilden.

Eine die Mathematik ergänzende formale Sprache liefert das Berechnungsmodell der Informatik (► Kap. 8), bei dem Verarbeitungsprozesse im Vordergrund stehen. Die Betonung der Prozesskomponente in den Modellen der Informatik macht diese besonders gut geeignet, Eigenschaften der Informationsverarbeitung formal zu beschreiben. Die kognitionswissenschaftliche Theoriebildung sollte daher beide genannten Aspekte berücksichtigen: Theorien über kognitive Strukturen und Prozesse sollten präzise und mithilfe einer formalen Sprache formuliert sein. Gleichzeitig sollte in den Theorien der erfahrungswissenschaftliche Gegenstand, also die Aspekte menschlicher Kognition, über die in einer Theorie Aussagen getroffen werden, im Vordergrund stehen. Eine empirische Prüfung der Validität solcher Theorien ist also unverzichtbar.

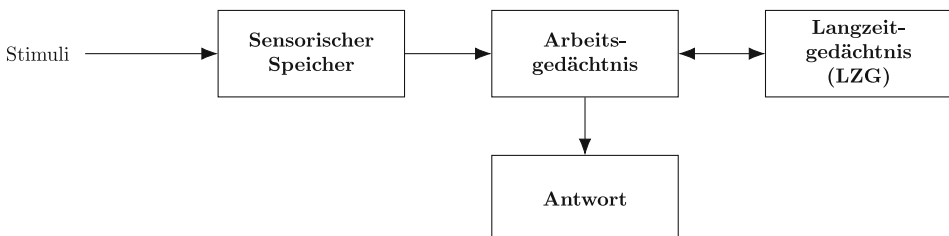
Im Folgenden werden wir die Bezeichnungen „Kognitionswissenschaft“ (engl. *cognitive science*) und „Kognitionsforschung“ synonym verwenden. Gehen wir speziell auf psychologische Beiträge zur Kognitionsforschung ein, so sprechen wir von „kognitiver Psychologie“ oder „Kognitionspsychologie“.

2.1 Der Informationsverarbeitungsansatz

Grundlage der meisten kognitionspsychologischen Theorien ist der Informationsverarbeitungsansatz, der den menschlichen Geist als ein System beschreibt, das Informationen ähnlich einem Computer in verschiedenen Stufen verarbeitet. Dieser Ansatz geht davon aus, dass Informationen aufgenommen, gespeichert, transformiert und abgerufen werden, was zentrale kognitive Prozesse wie Wahrnehmung, Gedächtnis und Problemlösen umfasst (etwa Lindsay und Norman 1977; Anderson 2000). Ein zentraler Ansatz, auf dem spätere Theorien aufbauen, ist das Multispeichermodell von Atkinson und Shiffrin (1968), das deutliche Parallelen zur Computerarchitektur² aufweist (■ Abb. 2.1).

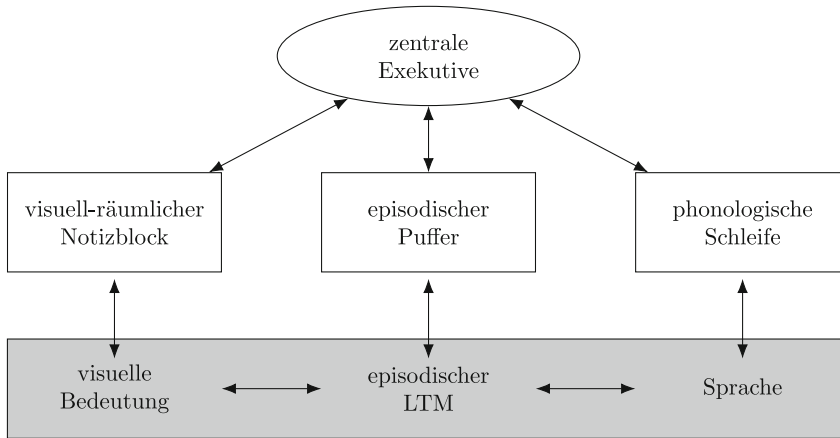
Hier werden eingehende Informationen (**Stimuli**) in aufeinanderfolgenden Stufen verarbeitet, wobei diese Stufen sich wechselseitig beeinflussen können. Umweltinformationen werden zunächst für einen Sekundenbruchteil im sinnesspezifischen **sensorischen Speicher** gehalten. Ein Teil dieser Informationen gelangt ins **Arbeitsgedächtnis** (AG), wo physikalische Reize in symbolische Informationen umgewandelt werden. Die im Arbeitsgedächtnis repräsentierten Informationen stehen für aktuelle kognitive Prozesse zur Verfügung, jedoch ist dessen Kapazität begrenzt. Nur eine begrenzte Menge an Information kann gleichzeitig aktiv verarbeitet und genutzt werden, was die Effizienz der Informationsverarbeitung beeinflusst. Die restlichen Informationen werden vergessen, da das sensorische Gedächtnis nur eine sehr flüchtige Speicherung ermöglicht.

Das zentrale Papier hierzu („Magical Number Seven“; Miller 1956) hat eine weitere Schlüsselfigur, George Miller, geschrieben. In diesem Artikel wurde einer der ersten systematischen Versuche zur Modellierung der kognitiven Ressourcen des Arbeitsgedächtnisses beschrieben und zu einer Codierung im Arbeitsgedächtnis, dem Chunking, was sich z. B. in der Encodierung von Zahlen durch Muster abbilden lässt. Die Telefonnummer 2 4 6 1 3 5 können Sie sich leicht merken, denn es sind die ersten geraden Zahlen und dann die ersten ungeraden Zahlen. Die Rolle des Kurzzeitgedächtnisses wurde, wie oben erwähnt, durch das Arbeitsgedächtnis als ein Prozessmodell anstelle eines „statischen Speichers“ abgelöst (Baddeley 2000). Dieses Modell ist noch viel näher an der Computermetapher (vgl. ■ Abb. 2.2), in diesem Modell existieren eine Zentrale Exekutive (vergleichbar mit der CPU eines Computers), welche im Arbeitsgedächtnis die



■ **Abb. 2.1** Die stufenartige Verarbeitung von Informationen erfolgt nach dem Modell von Atkinson und Shiffrin (1968) über den sensorischen Speicher, das Kurzzeit- und das Langzeitgedächtnis. In diesem Modell wurde das Kurzzeitgedächtnis durch das heutige vorherrschende Arbeitsgedächtnis ersetzt (s. u.)

² Mit „Computerarchitektur“ ist hier speziell die Architektur des sogenannten Von-Neumann-Rechners gemeint. John von Neumann schlug 1946 ein Konzept zur Gestaltung eines universellen Rechners vor, an dem sich die meisten modernen Computer orientieren.



■ **Abb. 2.2** Die aktuelle Version des Mehrkomponenten-Arbeitspeichermodells nach Baddeley (2000) umfasst die Zentrale Exekutive, den visuell-räumlichen Notizblock und die artikulatorische Schleife als Kernkomponenten des Arbeitsgedächtnisses. Baddeley ergänzte das Modell später um den episodischen Puffer. In grau sind Elemente des Langzeitgedächtnisses enthalten

Informationsverarbeitung steuert, und kleinere informationsspezifische Speicher, wie ein visuell-räumlicher Speicher und ein sprachbasierter Speicher, welcher zum Beispiel für das Merken von Zahlenfolgen eingesetzt wird, sowie ein episodischer Speicher, der visuelle und räumliche Informationen von Ereignissen inklusive von Erinnerungen aus dem Langzeitgedächtnis integrieren kann. Diese Aspekte könnten Sie an den RAM- oder Arbeitsspeicher eines Computers erinnern.

Zusätzlich zu Informationen, die über die Sinnesorgane ins Arbeitsgedächtnis gelangen, können dort auch aus dem **Langzeitgedächtnis (LZG)** aktivierte Informationen für die Verarbeitung bereitgehalten werden. Im Langzeitgedächtnis gespeicherte kognitive Strukturen wirken dabei auf die Kategorisierung und Wahrnehmung spezifischer sensorischer Reize ein. So sehen wir nicht alles, was das Auge wahrnimmt, sondern nur das, worauf wir unsere Aufmerksamkeit richten. Abhängig von aktuellen Zielen und weiteren Faktoren wird ein Teil der Informationen ins Langzeitgedächtnis übernommen (Lernen). Es wird angenommen, dass Wissen im Langzeitgedächtnis permanent gespeichert ist und Sie zum Beispiel an die Festplatte im Computer erinnern kann. „Vergessen“ ist dann auf ein Nichtauffinden von Information zurückzuführen, das nicht mit einem Löschen der Information gleichzusetzen ist. Zudem wird angenommen, dass Wissen im Langzeitgedächtnis nicht als unzusammenhängende Menge von Einzelinformationen, sondern strukturiert repräsentiert ist. Das Wissen im Langzeitgedächtnis wird häufig nach verschiedenen Aspekten unterschieden. Zum Beispiel unterscheidet Anderson (1996) zwischen der Repräsentation von Faktenwissen (deklaratives Wissen) und Regelwissen (prozedurales Wissen). In den folgenden Kapiteln werden verschiedene Konzeptionen für die Repräsentation von Wissen vorgestellt. Dabei liefern Annahmen und Erkenntnisse über die Charakteristika menschlicher Informationsverarbeitungsprozesse – wie zum Beispiel die beschränkte Kapazität des Arbeitsgedächtnisses – wichtige Randbedingungen für den Entwurf von Modellen.

2.2 Philosophische Grundlagen der Kognitionsforschung

In der Kognitionsforschung werden kognitive Prozesse als Informationsverarbeitungsprozesse beschrieben, was aber die Annahme eines grundlegenden erkenntnistheoretischen Paradigmas ist. Dieses Informationsverarbeitungsparadigma lässt sich wie folgt charakterisieren (siehe Stillings et al. 1998, zitiert nach Strube 1993, S. 304):

Annahme 1: Unterschiedliche kognitive Prozesse basieren auf gemeinsamen allgemeinen und grundlegenden Prinzipien.

Annahme 2: Kognition lässt sich unabhängig von ihrem materiellen Substrat, also der Neuronenaktivität, betrachten.

Annahme 3: Informationsverarbeitende Prozesse lassen sich als formale Prozesse beschreiben.

Annahme 4: Informationsverarbeitende Prozesse operieren auf Repräsentationen.

Eng damit verbunden ist die Annahme, dass sich dieses Informationsparadigma am besten durch einen Symbolverarbeitungsansatz modellieren lässt. Der oben bereits erwähnte KI-Forscher Allen Newell (1980) hat das in der „*physical symbol systems hypothesis*“ explizit formuliert (siehe auch Newell und Simon 1976).³ Die Hypothese besagt, dass alle kognitiven Prozesse nichts anderes als **Transformationen von physikalischen Symbolstrukturen** sind. Für unsere Betrachtungen genügt die abgeschwächte Hypothese, dass sich kognitive Prozesse als Transformationen von Symbolstrukturen **beschreiben** lassen.

Symbolstrukturen⁴ und Regeln für ihre Transformation liefern die Basis für eine einheitliche Beschreibung kognitiver Prozesse (► Annahme 1). Symbolstrukturen werden durch Verknüpfung elementarer Symbole aufgebaut, wobei die Verknüpfung syntaktischen Regeln gehorcht. Eine einfache Symbolstruktur wäre die folgende:

- Elementare Symbole: $S = \{0, 1, \#\}$
- Symbolstrukturen E können beispielsweise durch folgende syntaktische Regeln R_i aus Symbolen aus S aufgebaut werden:
 - $R_1: E \rightarrow \#$,
 - $R_2: E \rightarrow 0E$,
 - $R_3: E \rightarrow 1E$.

Dabei kann der Pfeil gelesen werden als „... wird ersetzt durch ...“. Beispielsweise wird durch die zweite Regel das Symbol E durch die beiden Symbole $0E$ ersetzt. Damit können längere Zeichenketten erzeugt werden. Die Anwendung dieser Regeln ermöglicht es zum Beispiel, folgende Strukturen zu erzeugen:

$E \rightarrow \#$ (Anwendung von R_1)
 $E \rightarrow 0E \rightarrow 0\#$ (Anwendung von R_2 und R_1)
 $E \rightarrow 1E \rightarrow 10E \rightarrow 100E \rightarrow 100\#$ (Anwendung von R_3 , R_2 , R_2 und R_1)

Damit ist E beschränkt auf beliebige (auch leere) Folgen aus den Ziffern 0 und 1, die mit dem Symbol $\#$ abgeschlossen sind. Das Symbol $\#$ kann als Zeichen für die leere Zeichenfolge/Symbolstruktur interpretiert werden. **Formale Sprachen**, die durch Symbolfolgen

3 In ► Kap. 12 werden wir auf die durch die „*Physical symbol systems*“-Hypothese ausgelöste philosophische Debatte eingehen.

4 In der theoretischen Informatik wird dies auch als Wort (endliche Symbolfolge) bezeichnet.

und syntaktische Regeln definiert sind, spielen nicht nur in der Theorie, sondern auch in der Praxis der Informatik eine zentrale Rolle. Im Kontext von Computern werden solche Sprachen verwendet, um alle Arten von Informationen zu verarbeiten. Da Computer nur mit elektrischen Zuständen arbeiten können, die zwei Zustände („an“ und „aus“) darstellen, erfolgt die Codierung von Informationen im sogenannten binären Zahlensystem, das nur die Symbole 0 und 1 kennt. Das bedeutet, dass wir, wenn wir beliebige Symbolstrukturen mit dem Computer verarbeiten wollen, eine Übersetzung aller Symbolstrukturen in binäre Symbolstrukturen benötigen. Wir demonstrieren im Folgenden, wie wir Dezimalzahlen auf binäre Zahlen abbilden können.

Im Dezimalsystem stehen uns die Ziffern 0 bis 9 zur Verfügung. Zahlen werden aus Summen von Zehnerpotenzen gebildet. So ergibt sich die Dezimalzahl 86 als

$$86_{10} = 8 \cdot 10^1 + 6 \cdot 10^0.$$

Die Einerstelle wird durch Multiplikation der Ziffer mit $1 = 10^0$ gebildet (für alle Zahlen x gilt $x^0 = 1$). Die Zehnerstelle wird durch Multiplikation der Ziffer mit $10^1 = 10$ gebildet, die Hunderterstelle durch Multiplikation der Ziffer mit $10^2 = 100$ und so fort. Im binären Zahlensystem haben wir nur die Ziffern 0 und 1 zur Verfügung. Hier werden Zahlen aus Summen von Zweierpotenzen gebildet. So können wir die Dezimalzahl 86 durch

$$\begin{array}{cccccccc} 1 \cdot 2^6 & + & 0 \cdot 2^5 & + & 1 \cdot 2^4 & + & 0 \cdot 2^3 & + & 1 \cdot 2^2 & + & 1 \cdot 2^1 & + & 0 \cdot 2^0 & = \\ 64 & + & 0 & + & 16 & + & 0 & + & 4 & + & 2 & + & 0 & = 86_{10} \end{array}$$

als die Binärzahl 1010110 repräsentieren. Auch beliebige Zeichenfolgen, wie zum Beispiel Worte wie „Vogel“ oder „Hund“ lassen sich durch solche 0–1-Kombinationen ausdrücken. Im Computer werden solche Begriffe auf diese Art, im binären Zahlensystem, als Folge von ASCII-Zeichen (*American Standard Code for Information Interchange*) codiert: „V“ hat beispielsweise die ASCII-Code-Nummer 86; im binären Zahlensystem wird daraus 1010110. Im Folgenden verzichten wir meist auf diese schwer lesbare Darstellung und verwenden Worte oder Dezimalzahlen als Elementarsymbole.

Bisher haben wir gezeigt, wie aus elementaren Symbolen mithilfe von syntaktischen Regeln komplexere Symbolstrukturen aufgebaut werden können. Wir haben eine **Grammatik** angegeben, die den syntaktischen Aufbau von Symbolstrukturen festlegt. Verwenden wir als Elementarsymbole Worte, wie zum Beispiel

$$S = \{\text{Peter, Eva, lacht, tanzt}\},$$

und als Regeln

$$\begin{array}{ll} R_1: & \text{Satz} \rightarrow \text{Eigennamen Verb}, \\ R_2: & \text{Eigennamen} \rightarrow \text{Peter}, \\ R_3: & \text{Eigennamen} \rightarrow \text{Eva}, \\ R_4: & \text{Verb} \rightarrow \text{lacht}, \\ R_5: & \text{Verb} \rightarrow \text{singt}, \end{array}$$

so können wir einfache Sätze wie „Peter lacht“ als Symbolstrukturen erzeugen. Um **Transformationen** von Symbolstrukturen zu beschreiben, benötigen wir weitere Regeln.

Diese Regeln müssen angeben, wie ein bestimmter Ausdruck E einer Symbolstruktur durch neue Ausdrücke ersetzt werden kann. Es könnte zum Beispiel eine Regel für die Addition definiert werden, die für Ausdrücke der Form $x + y$ die Addition ausführt. Für $E = 17 + 11$ würde die Regel also das Ergebnis 28 liefern. Die Symbolstruktur E wird durch Regelanwendung in eine neue Symbolstruktur, etwa $E' = 28$, transformiert. Regeln zur Symboltransformation liefern uns eine Möglichkeit, Informationsverarbeitungsprozesse wie etwa die Addition von Zahlen zu beschreiben. Umgangssprachlich formuliert könnte die Addition von zwei Zahlen, die aus je zwei Ziffern bestehen, durch folgende Regeln definiert werden:

- R_1 : Nimm von beiden Zahlen jeweils die hintere Ziffer und bilde ihre Summe und lösche die beiden Ziffern.
- R_2 : Notiere die letzte Ziffer der Summe als letzte Ziffer des Additionsergebnisses.
- R_3 : Besteht die Summe aus mehr als einer Ziffer, dann merke dir die noch nicht notierte Ziffer als Übertrag.
- R_4 : Nimm die vordere Ziffer der beiden Zahlen und, falls ein Übertrag existiert, den Übertrag, bilde ihre Summe und lösche die beiden Ziffern.
- R_5 : Notiere die letzte Ziffer der neuen Summe als vorletzte Ziffer des Additionsergebnisses.
- R_6 : Besteht die Summe aus mehr als einer Ziffer, dann notiere die noch nicht notierte Ziffer als vordere Ziffer des Additionsergebnisses.

In ► Kap. 8 und 9 werden wir eine spezielle Form von Transformationsregeln, sogenannte **Produktionsregeln**, einführen. Wie wir dort zeigen werden, eignen sich Produktionsregeln besonders gut zur Modellierung von Informationsverarbeitungsprozessen. Eine **Menge von Elementarsymbolen, Regeln zum Aufbau syntaktisch korrekter Symbolstrukturen** aus diesen Elementarsymbolen und **Regeln zu ihrer Transformation** definieren zusammen ein formales System zur Beschreibung von Informationsverarbeitungsprozessen, so wie es in Annahme 3 gefordert wird. Die Transformationsregeln operieren nicht auf konkreten Dingen der Welt, sondern auf Symbolstrukturen, die diese Dinge repräsentieren, so wie es in Annahme 4 gefordert wird.

Die zweite Annahme besagt, dass informationsverarbeitende Prozesse durch symbolische Codierungen beschrieben werden können, die unabhängig von der Materie (der „Hardware“) sind. In letzter Konsequenz heißt das auch, dass symbolische Codierungen Beschreibungen für die Informationsverarbeitungsprozesse verschiedener Systeme sind, egal ob diese Prozesse auf der „Maschine Gehirn“ oder dem Computer realisiert sind. Diese Auffassung wird nicht von allen Kognitionswissenschaftlern geteilt. Man kann durchaus moderatere Annahmen machen, wie zum Beispiel, dass Computersimulationen über Einzelaspekte menschlicher Denkprozesse Aufschluss geben können. Dabei ist es nicht notwendig, von einer Äquivalenz zwischen Computern und Menschen als symbolverarbeitende Systeme auszugehen, wohl aber von der Zulässigkeit, Informationsverarbeitungsprozesse unabhängig von den Prozessen des Gehirns zu beschreiben.

Die Annahme, dass geistige Prozesse unabhängig von ihrer materiellen Realisierung betrachtet werden können, liefert den Hintergrund für die **komputationale Theorie des Geistes**, eine wichtige Strömung der analytischen Philosophie. In dieser Theorie wird argumentiert, dass menschliche Denkprozesse auf **mental**en **Repräsentationen** operieren. Diese Repräsentationen können als syntaktisch strukturierte Symbole beschrieben wer-